Crystal structure of UDP-*N*-acetylglucosamine enolpyruvyltransferase, the target of the antibiotic fosfomycin

Ernst Schönbrunn[1]*, Stefan Sack[1], Susanne Eschenburg[2], Anastassis Perrakis[2†], Florian Krekel[3], Nikolaus Amrhein[3] and Eckhard Mandelkow[1]

Background: The ever increasing number of antibiotic resistant bacteria has fuelled interest in the development of new antibiotics and other antibacterial agents. The major structural element of the bacterial cell wall is the heteropolymer peptidoglycan and the enzymes of peptidoglycan biosynthesis are potential targets for antibacterial agents. One such enzyme is UDP-*N*-acetylglucosamine enolpyruvyltransferase (EPT) which catalyzes the first committed step in peptidoglycan biosynthesis: the transfer of the enolpyruvyl moiety of phosphoenolpyruvate (PEP) to the 3-hydroxyl of UDP-*N*-acetylglucosamine (UDPGlcNAc). EPT is of potential pharmaceutical interest because it is inhibited by the broad spectrum antibiotic fosfomycin.

Results: The crystal structure of substrate-free EPT has been determined at 2.0 Å resolution. The structure reveals a two-domain protein with an unusual fold (inside out α/β barrel) which is built up from the sixfold repetition of one folding unit. The only repetitive element in the amino acid sequence is a short motif, Leu-X_3-Gly(Ala), which is responsible for the formation of hydrogen-bond interactions between the folding units. An enzyme which catalyzes a similar reaction to EPT, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), has a very similar structure despite an amino acid sequence identity of only 25 %. To date, only these two enzymes appear to display this characteristic fold.

Conclusions: The present structure reflects the open conformation of the enzyme which is probably stabilized through two residues, a lysine and an arginine, located in the cleft between the domains. Binding of the negatively charged UDPGlcNAc to these residues could neutralize the repulsive force between the two domains, thereby allowing the movement of a catalytically active cysteine residue towards the cleft.

Addresses: [1]Max-Planck-Unit for Structural Molecular Biology, Notkestr. 85, c/o DESY, D-22603 Hamburg, Germany, [2]European Molecular Biology Laboratory, Notkestr. 85, c/o DESY, D-22603 Hamburg, Germany and [3]Institute of Plant Sciences, Swiss Federal Institute of Technology, Universitätsstr. 2, CH-8092 Zürich, Switzerland.

[†]Present address: Netherlands Cancer Institute (NKI), Department H2, Plesmanlaan 121, NL-1066 CX Amsterdam, Netherlands.

*Corresponding author.
E-mail: schoenbrunn@mpasmb.desy.de

Key words: domain movement, folding, hinge, peptidoglycan biosynthesis, sequence motif

Received: 7 June 1996
Revisions requested: 5 July 1996
Revisions received: 23 July
Accepted: 31 July

Structure 15 September 1996, 4:1065–1075

Introduction

In recent years much attention has been focused on enzymes that are involved in the biosynthesis of the bacterial cell wall [1]. The major structural element of the cell wall is the heteropolymer peptidoglycan. This macromolecule consists of alternating molecules of *N*-acetylglucosamine and *N*-acetylmuramic acid, cross-linked by peptide chains which are attached to the lactate group of the muramic acid residues. The initial step in peptidoglycan biosynthesis is the enolpyruvyl transfer from phosphoenolpyruvate (PEP) to the 3-hydroxyl group of UDP-*N*-acetylglucosamine (UDPGlcNAc), catalyzed by UDP-*N*-acetylglucosamine enolpyruvyltransferase (EPT; EC 2.5.1.7). The product of this reaction is the 3-enolpyruvyl ether of enolpyruvyl-UDPGlcNAc (EP-UDPGlcNAc). The formation of the muramyl sugar is completed by the reduction of this product to the 3-lactyl ether UDP-MurNAc, catalyzed by the flavoenzyme MurB. The free lactyl carboxyl group of UDP-MurNAc serves as the site of attachment for the peptide portion of the cell wall.

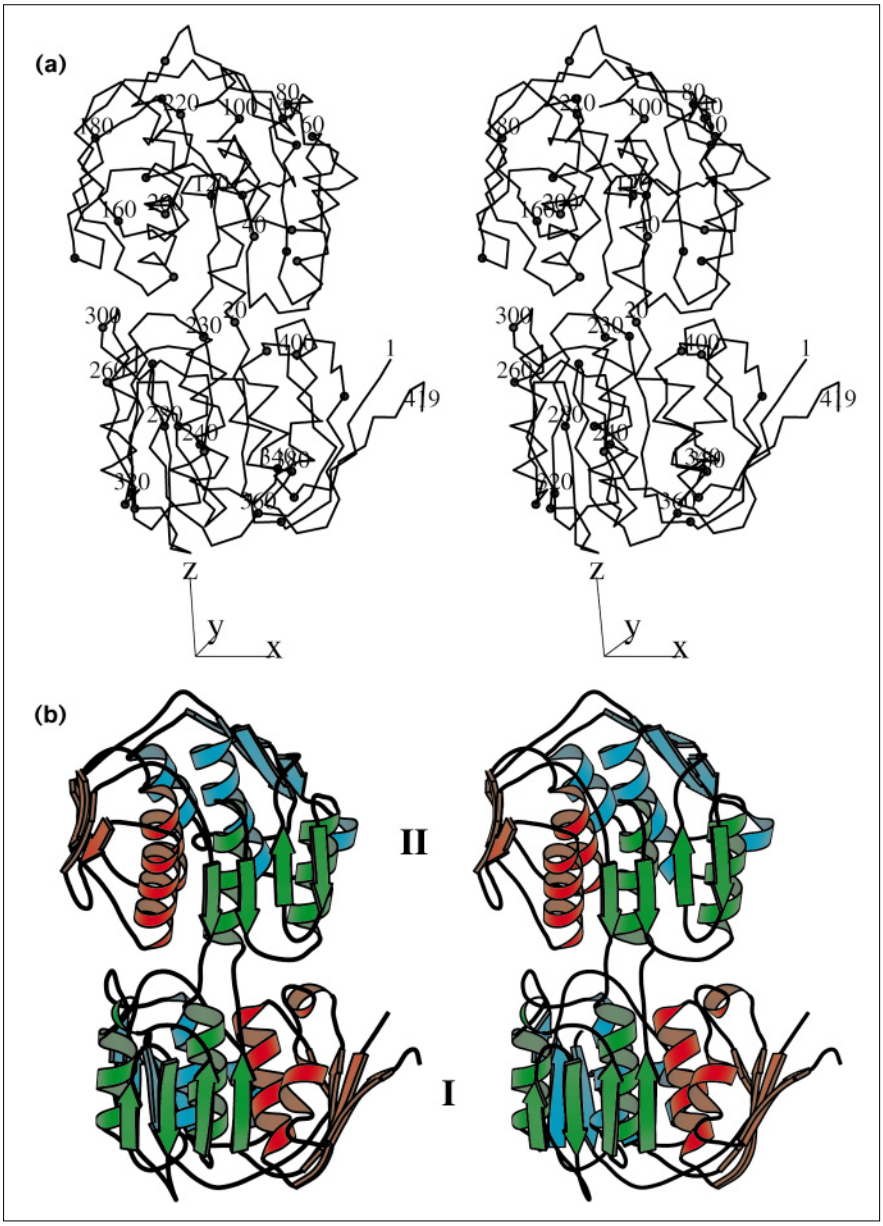
Historically, the pathway of peptidoglycan biosynthesis has been an important target for antibacterial agents [1]. The most widely used antibiotics are the β-lactams (e.g. penicillin) which inactivate the transpeptidases responsible for the cross-linking of individual polymers in the periplasmic space. EPT is of potential pharmaceutical interest because it is inhibited by the naturally occurring antibiotic fosfomycin ([1*R*,2*S*]-1,2-epoxypropyl phosphonic acid) [2]. For instance, fosfomycin trometamol salt is an orally administered antibiotic used in single-dose therapy of uncomplicated urinary tract infections [3]. Recently, the mechanism of action of fosfomycin on EPT has been re-evaluated using the large quantities of homogeneous enzyme available from the overexpression of *Enterobacter cloacae* EPT [4] and *Escherichia coli* EPT [5]. Both studies demonstrated that fosfomycin inactivates the enzyme by covalent modification of a cysteine residue (Cys115).

The increasing number of antibiotic resistant bacteria is a common clinical problem. This phenomenon has led to a

search for new strategies for the better understanding of the molecular basis of antibacterial agents. The elucidation of the three-dimensional structures of enzymes involved in crucial biochemical pathways unique to bacteria can be considered as an indispensable step towards the design of new antibiotics. For example, the structure of MurB has recently been determined in both its substrate-bound [6] and substrate-free [7] forms.

In most PEP-dependent enzymatic reactions PEP serves as a phosphoryl transfer agent. The reaction catalyzed by EPT is mechanistically unusual as it involves the attack at the electrophilic C2 position of PEP leading to the cleavage of the C–O bond [8]. The only other enzyme known in nature to catalyze the transfer of the intact enolpyruvyl moiety of PEP to a substrate is 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS; EC 2.5.1.19). EPSPS catalyzes the sixth step in the biosynthetic pathway of aromatic amino acids (the shikimate pathway) in bacteria, plants and fungi. EPSPS has been extensively studied as it is the target of glyphosate, the active ingredient of the broad-spectrum herbicide Roundup® [9]. Even though EPT and EPSPS share similarities in their amino acid sequences (25% identity and 47% similarity; [10]), their reaction mechanisms appear to be substantially different. For EPT, a covalent intermediate has been identified, in

Figure 1



The two-domain structure of EPT. (a) Stereo view of the Cα trace with every tenth residue marked and every twentieth residue labeled. (b) Stereo view of a ribbon plot of EPT. The individual folding units are displayed in red (Ia and IIa), green (Ib and IIb) and blue (Ic and IIc); this color-coding is used throughout the paper. An orthogonal coordinate system is chosen in such a manner that its origin is the center of mass (COM) of the molecule; the z axis pointing towards the top of domain II, and the y axis pointing away from the hinge. Throughout the text, the top and bottom domains are designated domains II and I, respectively. (The figure was drawn using MOLSCRIPT [35].)

which PEP is attached to the enzyme as the *O*-phosphothioketal of pyruvate through Cys115 [4,5]. In contrast, the EPSPS reaction proceeds through a tightly, but non-covalently bound tetrahedral intermediate [11,12]. Recent studies of the *Escherichia coli* EPT have provided evidence that the reaction pathway of EPT may involve the formation of both a covalently bound intermediate as well as a non-covalently bound tetrahedral intermediate [13].

Remarkably, EPSPS is not inactivated by fosfomycin, and conversely glyphosate does not inhibit EPT [14]. The crystal structure of substrate-free EPSPS has been determined at 3 Å resolution [15]. It is a two-domain structure with a unique fold that appears to be formed by a sixfold replication of one protein folding unit.

We report here the crystal structure of EPT in the absence of substrates and inhibitor at 2 Å resolution. The structure is very similar to that reported for EPSPS. However, the high resolution data for the EPT structure allow a much more accurate refinement of the atomic parameters and therefore contribute to a better understanding of structural and catalytic homologies between these two unique enzymes. In addition, we present the consequence of a repetitive motif found in the primary structure and suggest an important role of two residues with respect to catalysis.

Results and discussion

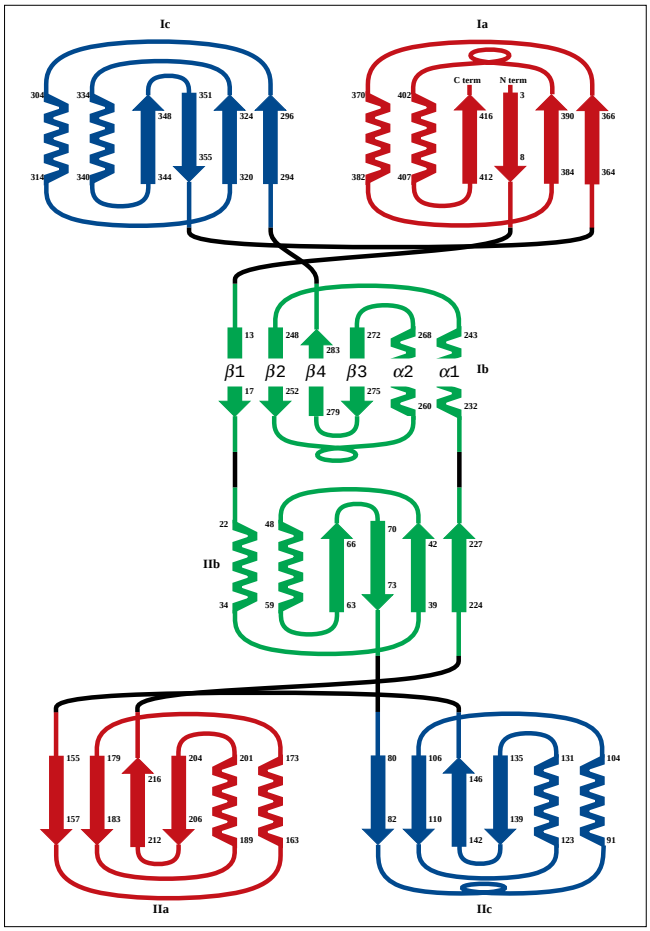
EPT was crystallized as previously described [16]. The electron-density map, based on the multiple isomorphous replacement with anomalous scattering (MIRAS) technique, was of excellent quality and allowed the complete polypeptide to be traced. Except for the last two C-terminal residues, a few side chains at the surface of the molecule, and a loop region around Cys115, there is continuous, well defined electron density for the entire polypeptide chain. Details of the structure determination and refinement are reported in the Materials and methods section.

Protein architecture

EPT consists of two globular domains (I and II) which are connected by a double-stranded hinge (Fig. 1). Each domain is roughly spherical with a radius of about 20 Å and consists of six helices and three four-stranded β sheets. 33% of the residues are located in α helices and 26% in β sheets. The N and C termini are located in domain I and build-up the antiparallel strands of one β sheet. Domain I consists mainly of the C- terminal portion of the sequence (residues 232–419 and 1–17) whereas domain II is composed mainly of the N- terminal portion (residues 22–227). The domains are linked by two strands each of four amino acids in length (residues 18–21 and 228–231).

The most striking feature of the EPT structure is the sixfold repetition of one folding unit or subdomain (comprising β1, α1, β2, α2, β3 and β4). As shown in Figure 2, each of the folding units contains a four-stranded β sheet with both parallel (β1 and β2) and antiparallel (β3 and β4) strands, and two parallel helices (α1 and α2). Three of these folding units (Ic, IIa and IIc) are formed from contiguous segments of the amino acid sequence. In contrast, each of the other three units (Ia, Ib and IIb) contain one β strand which is not adjacent in the primary structure; these β strands are formed by residues 3–8 (β4 of Ia = antiparallel), 13–17 (β1 of Ib = parallel) and 224–227 (β1 of IIb = parallel). Thereby five out of six antiparallel β strands are joined by hairpin loops. With the exception of folding unit IIa, where both antiparallel strands are connected by a five-residue loop (residues 207–211), the connection between the corresponding strands, in units IIc (140–141), Ic (349–350), Ib (276–278) and IIb (67–69), are two- and three-residue hairpin loops, respectively.

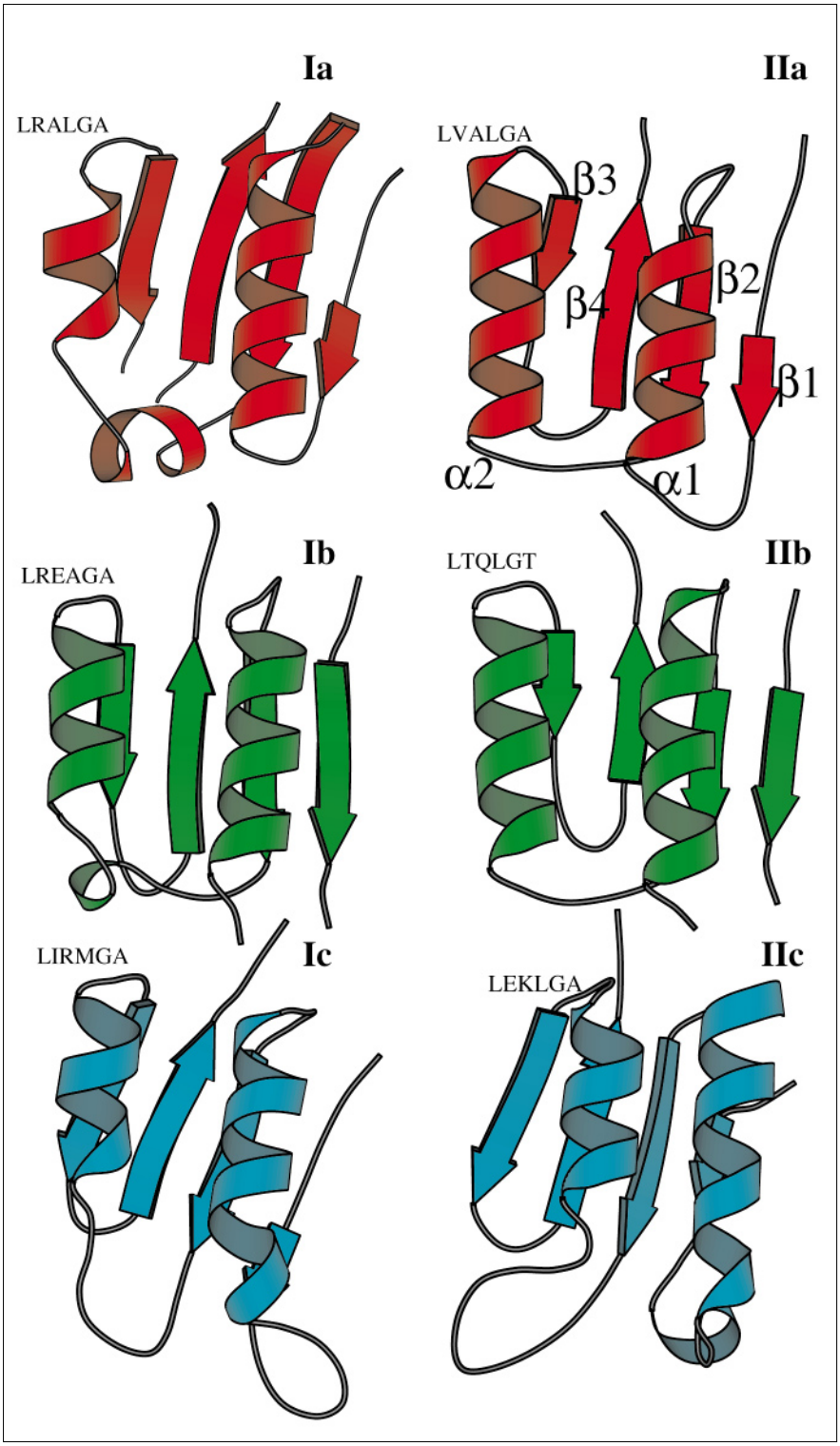
Figure 2



Topology sketch of EPT. The figure displays the sixfold repetition of one folding unit which consists of the elements β1, α1, β2, α2, β3 and β4. The folding units are designated Ia–Ic and IIa–IIc according to their location in the respective domains. The twisted loops in units Ia, Ib and IIc represent single-turn helices. The beginning and end of the secondary structure elements are labeled with the positions of the respective amino acids.

As shown in Figure 3, the α1 helices of the six folding units as well as the α2 helices of units IIa and IIb contain three to four turns, whereas the α2 helices of subdomains Ia, Ib, Ic and IIc are of the two-turn type. In addition to the α1 and α2 helices, there exist three single-turn helices. Two single-turn helices connect β2 with α2 (unit Ia, residues 393–396; and unit Ib, residues 256–258), and one connects β1 with α1 of unit IIc (residues 84–87).

Figure 3



Ribbon plots of the folding units of EPT as aligned through the α1 helix of unit Ia, using the program LSQMAN [36]. The occurrence of a repetitive amino acid sequence, $LX_3G(A)$, is shown; the sequence is part of a loop connecting the α2 helix with the β3 strand. The color scheme is the same as in Figure 1. (The figure was drawn with MOLSCRIPT [35].)

The cores of both domains consist of the three α1 helices which are surrounded by the three α2 helices, with solvent-accessible faces, and the three four-stranded β sheets (Fig. 4). One therefore might characterize the overall fold as an inside out α/β barrel.

The center of mass (COM) of EPT resides in the cleft between domain I and domain II. In an orthogonal coordinate system, with the COM as the origin (Fig. 1), the z axis can be considered as an approximate threefold symmetry axis relating the folding units within each domain. This becomes evident from the view through the domains in the direction of the hinge (Fig. 4). Additionally, the y axis relates the two domains by an approximate dyad (Fig. 1).

The repetitive motif Leu-X_3-Gly(Ala)

Although the folding units are remarkably similar in secondary structure elements and fold, the only repetitive element in the amino acid sequence is a short motif Leu-X_3-Gly(Ala) ($LX_3G(A)$), which is part of a loop connecting the α2 helix with the β3 strand (Fig. 3). This motif, first recognized by Marquardt *et al.* [17], is now shown to be responsible for the attachment of the folding units to each other (Fig. 4). As displayed in Table 1, the carbonyl oxygen of every fourth residue of the motif is involved in establishing a hydrogen bond to a backbone amide of the adjacent folding unit. The side chain of the conserved leucine residue always points towards the interior. This motif forms a hydrophobic pocket together with the α1 helix of the same folding unit, and the α1 helix of the adjacent unit (Fig. 4). Thereby, three hydrophobic pockets are formed per domain.

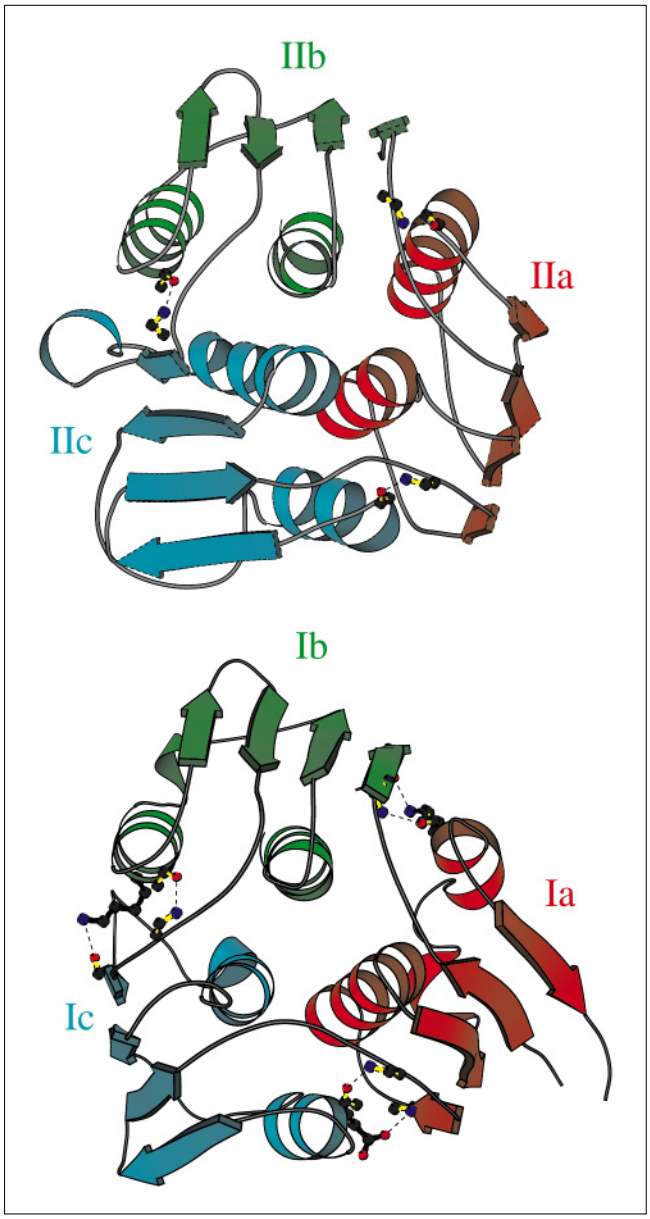
Apparently, the units of domain II are linked through backbone hydrogen bonds only, whereas in domain I there are additional hydrogen bonds: between the ε-amino groups of two lysine residues (Lys405 and Lys265) and the carbonyl oxygens of two threonine residues (Thr17 and Thr293); and between the carboxyl oxygen of Glu337 and the backbone amide of Met366 (Fig. 5). The two lysines and the glutamic acid residue are in position n–1, where n is the conserved leucine. The superposition of all $LX_3G(A)$ regions reveals the almost identical architecture of this motif throughout the structure with root mean square (rms) distances between 0.14 Å and 0.26 Å.

The hinge region

The two strands connecting both domains display excellent electron density suggesting a high degree of rigidity. Both strands are connected to each other by two backbone hydrogen bonds. The residues involved in hydrogen bonding are Asp231-NH and Ser19-CO (distance = 2.9 Å), and Leu229-CO and Ala21-NH (distance = 2.9 Å). One additional connection between the two domains is a salt bridge located in close proximity to the cross-over strands; the bridge is formed by the carboxyl oxygen of Glu190 (in domain II) and the guanidinium group of Arg232 (in domain I).

Several tightly bound water molecules are observable in this region and probably serve as stabilizers. Both strands

Figure 4



Ribbon plots of the two domains of EPT, as aligned through the α1 helix of folding unit Ib using the program LSQMAN [36]. The orientation is chosen to show the view through the cores of both domains in the direction of the hinge region; from the top (–z direction, domain II) and bottom (+z direction, domain I) of the molecule, respectively. The $LX_3G(A)$-dependent main-chain and side-chain atoms responsible for the formation of hydrogen bonds connecting the folding units to each other, are shown in ball-and-stick representation. The color scheme is the same as in Figure 1. (The figure was drawn using MOLSCRIPT [35].)

Table 1

The $LX_3G(A)$-directed hydrogen bonds between the folding units.

Main chain–main chain interaction			Distance (Å)* Monomer A	Monomer B
Ia–Ib	LRAL409 GA	[C = O–HN] Q13	3.1	2.7
Ic–Ia	LIRM341 GA	[C = O–HN] S361	2.5	3.0
Ib–Ic	LREA269 GA	[C = O–HN] K290	2.7	3.1
IIa–IIb	LVAL201 GA	[C = O–HN] G222	3.2	3.0
IIc–IIa	LEKL132 GA	[C = O–HN] K152	2.9	2.6
IIb–IIc	LTQL60 GT	[C = O–HN] N78/N79	2.8/3.0	2.9/3.0
Side chain–main chain interaction				
Ia–Ib	K405 LRALGA	[Nζ – O=C] T17	2.7	2.7
Ic–Ia	E337 LIRMGA	[Oε1 – HN] M366	2.8	3.2
Ic–Ia	K265 LREAGA	[Nζ – O=C] T293	2.9	2.7

*The distances are given for both independently treated monomers, A and B, in the asymmetric unit.

have an equal length with distances between the Cα atoms of residues 18 to 21 and 228 to 231 of 9 Å.

Organization of the asymmetric unit

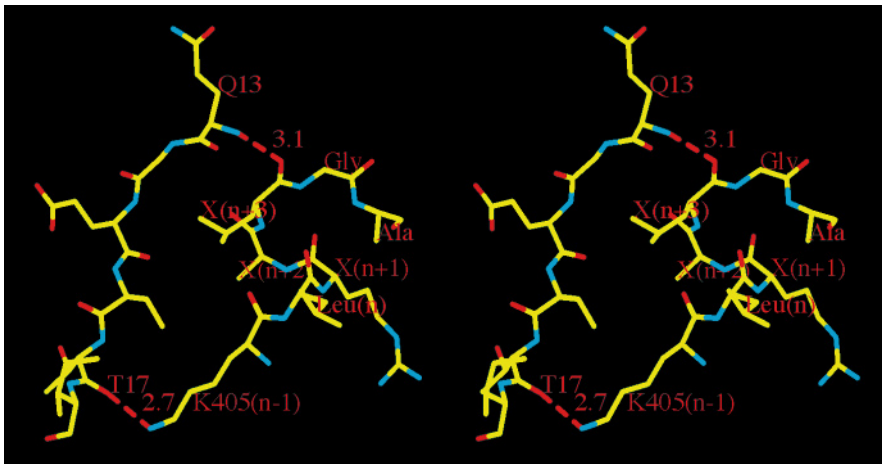
The crystals of EPT were grown in the presence of cyclohexylammonium ions, as described previously [16]. In the absence of this agent, only needles not useful for crystallographic analysis are formed. The asymmetric unit consists of two monomers, A and B. The interpretation of the electron-density map revealed that EPT dimerizes due to the interaction between Trp279 of both monomers, with two cyclohexylammonium ions creating a hydrophobic environment through the cyclohexyl moiety. The distance between the two indole moieties is 7.8 Å. The cyclohexyl moieties are located approximately midway between the two tryptophans. The transformation of monomer A to monomer B is given by a Eulerian rotation around the Crowther angles $\alpha = 178.2°$, $\beta = -64.6°$, $\gamma = 1.0°$ and a translation of $x = 10.4$ Å, $y = -43.9$ Å and $z = 8.0$ Å, referring to the coordinate system in Figure 1.

The PEP/fosfomycin-binding site

As mentioned above, Cys115 is involved in binding of PEP and the inhibitor fosfomycin. This active cysteine residue is located in a highly flexible loop comprising 12 amino acids (residues 111–122) connecting β2 with α2 of subdomain IIc (Fig. 3). Although the electron density is not well defined in this region it is obvious that Cys115 is highly solvent accessible. This large loop is mounted through two proline residues (Pro112 and Pro121) suggesting a rope held by two anchors. Within a sphere of 9 Å around Cys115 there are no side chains apart from those building up the loop. The closest neighbours are Glu140 (distance ≈ 10 Å) and Arg91 (distance ≈ 13 Å). As the present structure reflects the enzyme in the absence of substrate or inhibitor, it is reasonable to assume that the flexible loop undergoes

Figure 5

A stereo drawing of the region around the $LX_3G(A)$ motif of folding unit Ia. The hydrogen-bond interactions between unit Ia (Leu409 and Lys405) and unit Ib (Gln13 and Thr17) are shown. The conserved residues of the $LX_3G(A)$ motif are labeled Leu, Gly, and Ala, respectively; non-conserved residues are labeled with X's. Nitrogen atoms are depicted in blue, oxygen atoms in red. (The figure was produced with O [26].)



substantial structural changes upon binding of UDPGlcNAc and PEP or fosfomycin. The electron density around the loop region is different for both monomers in the asymmetric unit, again indicating the high degree of flexibility.

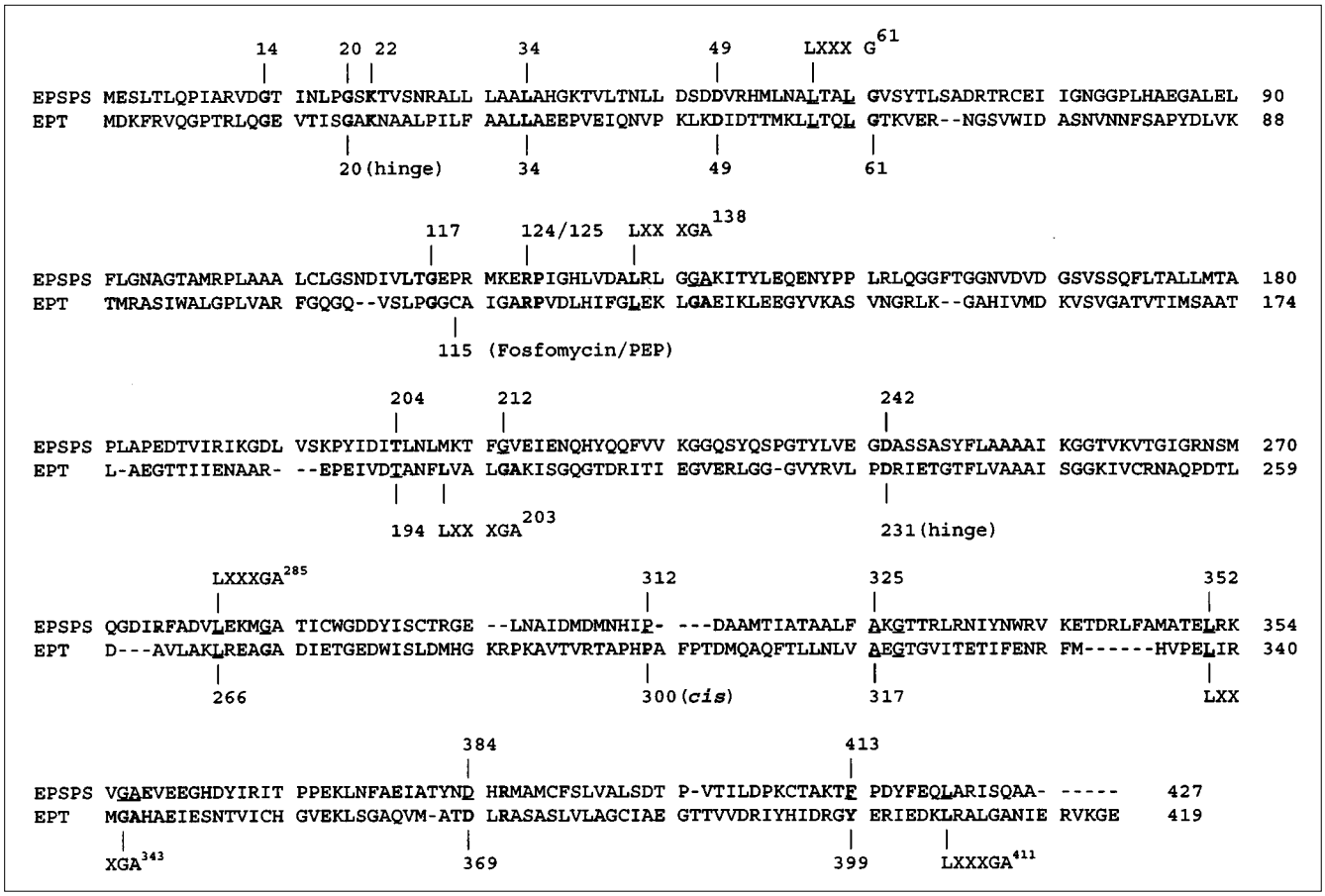
Comparison of the 3D structures of EPT and EPSPS

Like EPT, EPSPS is a two-domain protein in which both domains are connected by a double-stranded hinge [15]. The structural principle of the folding units is the same for both enzymes (Figs 2,3). Even the corresponding folding units containing the β strands formed from discontinuous segments of both polypeptides (units Ia, Ib, IIb) are identical to each other. In contrast to EPT, folding unit Ia of EPSPS is incomplete due to the missing β3 strand built up by the C terminus.

A precise structural comparison between the two enzymes is not possible because information concerning the side chain arrangement in EPSPS is missing. However, as the overall structure of both enzymes is almost identical it is reasonable to assume that several conserved residues adopt similar conformations in EPT and EPSPS.

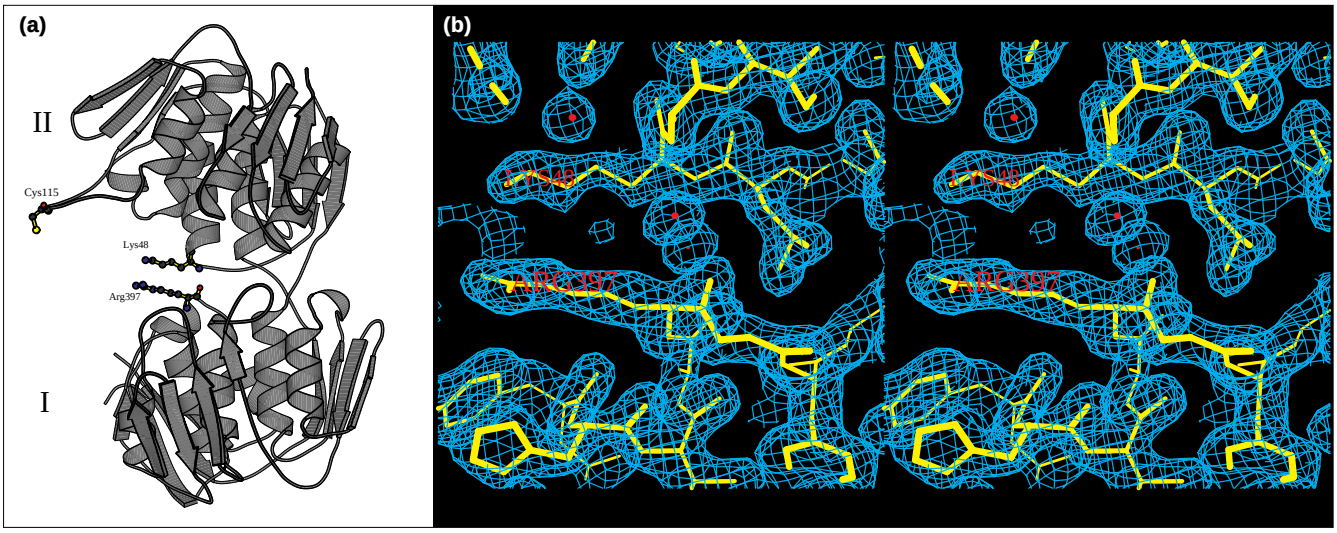
The sequence alignment (using CLUSTAL W [18]) of 20 EPSPS and six EPT primary structures, reveals 12 absolutely conserved residues and 17 residues being

Figure 6



The aligned amino acid sequences of *E. coli* EPSPS and *E. cloacae* EPT. The amino acid sequences of of 20 EPSPS's: *AROA_BACSU* (*Bacillus subtilis*); *AROA_LACLA* (*Lactococcus lactis*); *AROA_PETHY* (*Petunia hybrida*); *ARO1_TOBAC* (*Nicotiana tabacum*); *AROA_LYCES* (*Lycopersicon esculentum*); *AROA_ARATH* (*Arabidopsis thaliana*); *AROA_BRANA* (*Brassica napus*); *AROA_SALGL* (*Salmonella gallinarum*); *AROA_SALTI* (*Salmonella typhi*); *AROA_SALTY* (*Salmonella typhimurium*); *AROA_ECOLI* (*E. coli*); *AROA_KLEPN* (*Klebsiella pneumoniae*); *AROA_YEREN* (*Yersinia enterocolitica*); *AROA_HAEIN* (*Haemophilus influenzae*); *AROA_PASMU* (*Pasteurella multocida*); *AROA_AERSA* (*Aeromonas salmonicida*); *AROA_BORPE* (*Bordetella pertusis*); *AROA_BURPS* (*Burkholderia pseudomallai*); *AROA_STAAU* (*Staphylococcus aureus*); *AROA_MYCTU* (*Mycobacterium tuberculosis*); and six EPT's *MURA_ECOLI* (*E. coli*); *MURA_ENTCL* (*Enterobacter cloacae*); *MURA_HAEIN* (*Haemophilus influenzae*); *MURA_ACICA* (*Acinetobacter calcoaceticus*); *MURA_BACSU* (*Bacillus subtilis*); *MURZ_MYCLE* (*Mycobacterium leprae*) (as deposited in the Swiss Prot database), were aligned using the program CLUSTAL W [18]. Amino acid pairs marked in bold are absolutely conserved in all EPSPS's and EPT's. Underlined amino acids are conserved in more than 90 % of the respective sequences. The locations of the $LX_3G(A)$ motifs and fosfomycin/PEP-binding site are indicated.

Figure 7



Stabilization of the open conformation of EPT. (a) Ribbon plot of EPT illustrating the likely role of residues Lys48 and Arg397 with respect to catalysis and maintenance of the open conformation. The molecule is rotated 115° around the vertical axis (z axis), with respect to the orientation in Figure 1. The binding of UDPGlcNAc to these residues could neutralize the repulsive force between their positively charged ε-amino and guanidinium groups, thereby allowing the movement of domain II, or parts of domain II, with the active Cys115 moving towards the cleft. (The figure was drawn with MOLSCRIPT [35].) (b) A $2F_o–F_c$ stereo map contoured at 1σ displaying the parallel arrangement of Lys48 and Arg397 in the cleft. The distance between the ε-amino and the guanidinium groups is approximately 3.6 Å. The orientation is the same as in (a).

conserved in more than 90% of the sequences (Fig. 6). All of these conserved residues are well defined in the EPT structure, with the exception of the region around Cys115. For example, in EPT the absolutely conserved residues Lys22 and Asp49 are linked to each other by a salt bridge, between the respective carboxyl oxygen and ε-amino group (distance 2.8 Å). The absolutely conserved Leu34 forms a hydrophobic pocket together with Leu175, Leu198 and Leu221, thereby connecting folding unit IIa with unit IIb. Pro300 is responsible for the formation of a *cis*-peptide and part of the motif $PHP^{300}AFP$. This region is located close to the cross-over strands in the loop connecting β1 with α1 in folding unit Ic (Figs 2,3). Several well defined water molecules are in close proximity to the *cis*-peptide. The density of all residues around this motif is exceptionally well defined with side chain B factors between 10 $Å^2$ and 20 $Å^2$; this is indicative of a high degree of rigidity. The carboxyl oxygen of the conserved residue Asp369 forms a salt bridge with the guanidinium group of Arg331. Thus, folding units Ia and Ic are connected to each other: by two $LX_3G(A)$ dependent side chain–main chain hydrogen bonds; by one main chain–main chain hydrogen bond (Table 1); and through the salt bridge described above.

The $LX_3G(A)$ motif is also found in the EPSPS sequence with the exception of the third and sixth motif (EPSPS residues 205–212 and 420–426). A similar role of this motif, as shown here for the EPT structure, was not described for the EPSPS structure.

Implications for catalysis

In contrast to the binding site of shikimate-3-phosphate (S3P) in EPSPS, which is considered to reside in the cleft between the two domains [19], little is known about the residues of EPT involved in UDPGlcNAc binding. However, the high structural similarity between both enzymes suggests that binding of UDPGlcNAc by EPT also takes place in the cleft.

The two-domain architecture of EPT suggests large conformational changes mediated through the hinge region [20]. This is consistent with small angle X-ray scattering experiments where a significant decrease in the radius of gyration of EPT was observed upon binding of UDPGlcNAc [21]. The experimental radius of gyration of substrate-free EPT (R_g = 24.2 Å) was in good agreement with the theoretical value derived from the EPT coordinates (R_{gt} = 23.5 Å) using the program CRYSOL [22].

We propose that the structure of EPT shown here represents the open conformation which is probably stabilized through two residues in the cleft (Arg397 and Lys48) lying nearly parallel to each other (Fig. 7). The repulsive force between the positive charges of the ε-amino group and the guanidinium group might assure the open state of the substrate-free enzyme. As the sugar nucleotide (UDPGlcNAc) is negatively charged, because of its diphospho moiety, it could be attracted by these positive charges. Moreover, because the binding of substrates follows an ordered

mechanism with UDPGlcNAc binding first [23], we assume that binding of UDPGlcNAc neutralizes the repulsion between the two domains thereby allowing the movement of the active Cys115 towards the cleft. This motion could be accomplished by a movement of the whole of domain II or by the movement of specific folding units. As shown above (Fig. 4), the movement of parts of domain II rather than parts of domain I is plausible because the folding units in domain I are stabilized through the additional side chain–main chain hydrogen bonds (Table 1). It is reasonable to assume that the binding of UDPGlcNAc to EPT results in local changes of structural elements in the proximity of the cross-over strands, triggering in turn large conformational changes of the globular domains.

A comparable arrangement of side chains, as formed by Lys48 and Arg397 in EPT, is not evident in the EPSPS structure [15]. Instead, it was suggested that in EPSPS a helical macrodipole effect results in an accumulation of positive charges between the domains, which could be responsible for both the binding of substrate(s) and maintenance of the open conformation. In principle, this explanation is also valid for EPT, as the spatial orientation of the helices in both its domains is very similar to that in EPSPS.

Clearly, kinetic and structural studies on enzyme forms with site-directed mutations are required to confirm the suggested role of the respective residues in catalysis and maintenance of the structure. For evidence of the proposed large conformational changes upon catalysis, which are most probably mediated through the hinge region, we must await the structure determination of the EPT–UDPGlcNAc–fosfomycin complex.

Biological implications

UDP-*N*-acetylglucosamine enolpyruvyltransferase (EPT) is a key enzyme in the biosynthesis of peptidoglycan, the major structural component of the bacterial cell wall. The enzyme catalyzes the enolpyruvyl transfer from phosphoenolpyruvate (PEP) to a second substrate; the only other enzyme known in nature to catalyze this reaction is 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). Despite their catalytic similarities, the two enzymes differ substantially with respect to the structures of their second substrates, that is, UDP-*N*-acetylglucosamine (UDPGlcNAc) and shikimate-3-phosphate (S3P). Moreover, EPT and EPSPS are selectively inhibited by the antibiotic fosfomycin and by the herbicide glyphosate, respectively.

The crystal structure of substrate-free EPT reveals a surprisingly high structural similarity to the previously determined EPSPS structure. EPT is a two-domain protein which is built up by a sixfold repetition of one folding unit. Despite the structural similarity of these folding units, the only repetitive element of the amino acid sequence is a small motif, Leu-X_3-Gly(Ala), which is responsible for the attachment of the folding units to each other through hydrogen-bond interactions.

The only residue known to be involved in catalysis is Cys115, which is responsible for both the binding of PEP and fosfomycin. The region around Cys115 is a large flexible loop which presumably undergoes substantial structural changes upon catalysis. The binding of substrates follows an ordered mechanism with UDPGlcNAc binding first. We assume that binding of this sugar nucleotide by EPT takes place in the cleft between the two domains. This could result in large conformational changes, mediated by a hinge region, allowing the movement of the active Cys115 towards the cleft.

Further studies on EPT based on the structural information presented here will eventually lead to the development of a new generation of antibacterial agents. Their rational design may help to overcome the increasing problem of antibiotic resistance.

Materials and methods

Crystallization

EPT was crystallized as previously described [16]. Briefly, diffraction quality crystals were grown from 0.8 M sodium/potassium phosphate buffer (pH 6.4) containing 40 mM cyclohexylammonium phosphate. The crystals belong to space group C2 with $a = 86.9$ Å, $b = 155.9$ Å, $c = 83.8$ Å, $\beta = 91.6°$, and contain two molecules per asymmetric unit. The solvent content is 63 %.

X-ray data collection

All X-ray data were collected at the EMBL BW7b and X31 beamlines (DESY, Hamburg) equipped with MarResearch image plate scanners. Crystals were frozen in an evaporating stream of dry nitrogen (Oxford Cryosystems). The cryoprotectant consisted of 1.5 M sodium/potassium phosphate buffer (pH 6.4) containing 40 mM cyclohexylammonium phosphate and 35 % glycerol. Native and heavy-atom derivative data were collected from single crystals and were processed using the DENZO-SCALEPACK package [24] (Table 2).

Phasing

The crystal structure of EPT was determined by MIRAS phasing techniques using two heavy-atom derivatives. We did not perform molecular replacement (MR) using EPSPS as a model because only the Cα coordinates of EPSPS are available to the public. Native crystals were soaked in 5 mM $K_2Pt(CN)_4$ for 48 h or in 5 mM $KAu(CN)_2$ for 24 h. Data were collected at EMBL X31 beamline at a wavelength of 1.0 Å to optimize anomalous scattering. Both derivatives yielded clear anomalous signals (Table 2). The Patterson functions were solved for both derivatives using the program package PHASES [25], giving the positions of five gold and six platinum sites. Three of the gold sites were identical to the platinum sites. Heavy-atom positions and occupancies were refined with PHASES. All reflections between 25 and 2.0 Å resolution were included in the phasing. The overall figure of merit was then 0.76. The 2.0 Å MIRAS phases were improved through solvent flattening and the resulting electron-density map was skeletonized using GMAP as implemented in PHASES. It was judged that non-crystallographic symmetry (NCS) averaging was not necessary to proceed with model building due to the excellent quality of the map. The map was readily interpretable and a Cα trace was built into the density using the program O [26]. There was unambiguous density for most of the side chains and all of the 419 residues of one of the two monomers were placed using

Table 2

Statistics of data collection and phasing.

Data set	Platinum	Gold	Native
Resolution range (Å)	30–2	25–2	25–1.8
Total observations	657 859	702 087	744 941
Unique reflections	74 497	74 025	103 189
Completeness (%)	98.0 (95.8)†	97.9 (95.3)†	99.8 (98.7)†
R_{sym}* (%)	3.7 (13.6)†	4.5 (27.4)†	5.3 (29.0)†
$I/\sigma(I) > 2$ (%)	89.4 (74.8)†	85.0 (56.7)†	84.5 (60.9)†
Anomalous pairs	68 514	66 415	
Completeness of			
	anomalous pairs (%)	97.0 (92.1)†	96.9 (90.5)†
R_{sym}* (anomalous) (%)	3.0 (11.6)†	3.6 (23.8)†	
$I/\sigma(I) > 2$ (anomalous) (%)	86.4 (67.2)†	81.2 (48.4)†	
Number of sites	6	5‡	
Phasing power§			
	anomalous pairs	1.99 (1.93)†	1.41 (1.22)†
	isomorphous pairs	1.78 (1.12)†	1.66 (1.05)†
R_{cullis}††	0.65 (949)	0.69 (791)	
R_{kraut}# (anomalous pairs)	0.066 (55 601)	0.115 (42 835)	
R_{kraut}# (isomorphous pairs)	0.068 (109 008)	0.121 (83 730)	
Figure of merit**		0.605/0.761¶	
Number of phased reflections		57 441	

*$R_{sym} = \Sigma_{i,hkl}|I(i,hkl) - \langle I(hkl)\rangle|/\Sigma_{i,hkl}\langle I(hkl)\rangle$. †The numbers given in parentheses denote the respective values of the highest resolution shell. ‡Three of the five Au sites are equivalent to Pt.

§Phasing power, $[\Sigma n|F_H|^2/\Sigma n\ |E|^2]^{1/2}$ with $\Sigma n|E|^2 = \Sigma n\{|F_{PH}|(obs) - |F_{PH}|(obs)\}^2$.

††$R_{cullis} = \Sigma hkl||F_{PH} \pm F_P| - F_H(calc)|/\Sigma hkl|F_{PH} \pm F_P|$; the values in parentheses are the number of centric reflections.

#$R_{kraut} = \Sigma hkl||F_{PH}| - |F_P + F_H(calc)||/\Sigma hkl|F_{PH}|$; the values in parentheses are the number of acentric reflections. **Figure of merit, $m = |F(hkl)_{best}|/|F(hkl)|$, with $F(hkl)_{best} = \Sigma\alpha P(\alpha)F_{hkl}(\alpha)/\Sigma\alpha P(\alpha)$; ¶the figure of merit after solvent-flattening and phase extension.

the best fitting rotamers generated by O. The second monomer was located using the program AMoRe [27,28] and the corresponding model was then produced by applying the rotation-translation matrix on the chain of the first monomer.

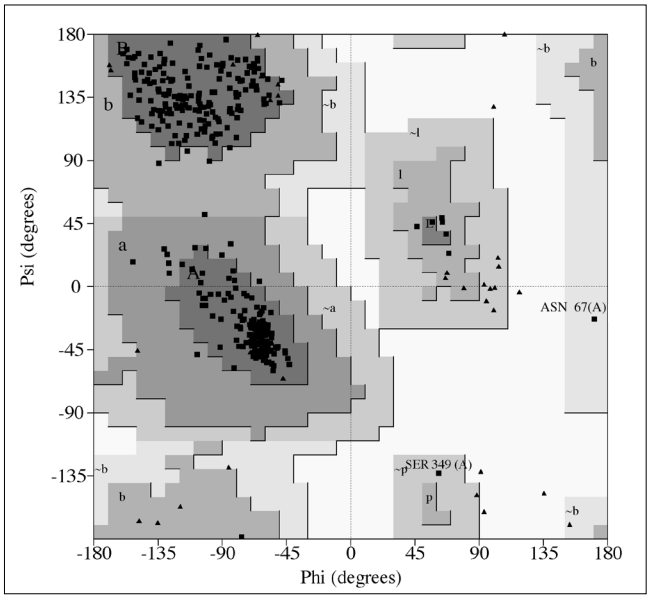
Structure refinement

The initial model was submitted to rigid-body refinement with X-PLOR [29] to optimize the exact orientation and position of each molecule in the asymmetric unit. Then, two cycles of simulated annealing (3000 K slowly cooled to 300 K) [30] were performed with X-PLOR followed by individual isotropic temperature factor refinement. The initial R factor of 52 % dropped to 31 % for data between 10 Å and 2 Å. The free R test [31] was applied throughout the refinement leaving 10 % of the data as a test set. The free R value at that stage converged to 37 %. At this stage water molecules were included using the program ARP [32]. Only water molecules with clear density in the $3F_o–2F_c$ map (peaks above 1σ) and the $F_o–F_c$ map (peaks above 3σ) and making plausible hydrogen-bond interaction were kept. A total of 476 water molecules were located and their positions refined with individual isotropic temperature factors. Further refinement was performed using the least-squares method as implemented in PROLSQ [33]. Both molecules were treated independently and after each cycle a $3F_o–2F_c$ and $F_o–F_c$ map was visually inspected using the program O and the model was corrected for errors. The current crystallographic R factor is 19.7 % (converging R_{free} = 27 %). Further refinement is in progress.

Quality of the model

On the basis of the stereochemical analysis with PROCHECK [34], 92.9 % of all residues of both monomers have dihedral angles within the most favored regions of the Ramachandran plot and 6.4 % within additionally allowed regions (Fig. 8). Ala119 of monomer B, Asn67 of monomer A, and Ser349 of both monomers reside in generously

Figure 8



Ramachandran plot for the EPT dimer. The plot was produced for 2×419 residues using PROCHECK [34]. The disallowed regions are shaded with the lightest gray. Glycine residues are indicated as triangles, and all other residues as squares.

allowed regions whereas Asn67 of monomer B falls within the disallowed region. Ala119 is part of a large flexible loop and Ser349 and Asn67 are members of β turns in folding units Ic and IIb, respectively. In contrast to the well defined region around Ser349, the polypeptide chain between Asn67 and Gly68 is interrupted in both monomers probably accounting for the improper main-chain angles. The rms deviations from ideality on bond lengths and bond angles are 0.016 Å and 2.0°, respectively. The average B factors are 20.1 Å^2 for main-chain atoms (in total 3352), 25.5 Å^2 for side-chain atoms (in total 2932), 29.9 Å^2 for water molecules (in total 476) and 34.2 Å^2 for the cyclohexylammonium atoms (in total 14).

Accession numbers

The atomic coordinates of the *E. cloacae* EPT have been deposited in the Brookhaven Protein Data Bank with accession code 1naw.

Acknowledgements

We wish to thank the staff of the EMBL outstation Hamburg for support. This study contains part of the doctoral thesis of SS who receives a grant from the Friedrich-Ebert-Foundation.

References

1. Bugg, T.D.H. & Walsh, C.T. (1993). Intracellular steps of bacterial cell wall peptidoglycan biosynthesis: enzymology, antibiotics and antibiotic resistance. *Nat. Prod. Rep.* **9**, 199–215.
2. Kahan, F.M., Kahan, J.S., Cassidy, P.J. & Kropp, H. (1974). The mechanism of action of fosfomycin (Phosphonomycin). *Ann. NY Acad. Sci.* **235**, 364–385.
3. Scaglione, F., Cicchetti, F., Demartini, G. & Arcidiacono, M. (1994). Fosfomycin distribution in the lower urinary tract after administration of fosfomycin trometamol salt. *Int. J. Clin. Pharmacol. Res.* **14**, 107–109.
4. Wanke, C. & Amrhein, N. (1993). Evidence that the reaction of the UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase proceeds through the *O*-phosphothioketal of pyruvic acid bound to Cys115. *Eur. J. Biochem.* **218**, 861–870.
5. Marquardt, J.L., *et al*., & Walsh, C.T. (1994). Kinetics, stoichiometry, and identification of the reactive thiolate in the inactivation of UDP-GlcNAc enolpyruvoyl transferase by the antibiotic fosfomycin. *Biochemistry* **33**, 10646–10651.
6. Benson, T.E., Filman, D.J., Walsh, C.T. & Hogle, J.M. (1995). An enzyme–substrate complex involved in bacterial cell wall biosynthesis. *Nat. Struct. Biol.* **2**, 644–653.
7. Benson, T.E., Walsh, C.T. & Hogle, J.M. (1996). The structure of the substrate-free form of MurB, an essential enzyme for the synthesis of bacterial cell walls. *Structure* **4**, 47–54.
8. Walsh, C.T., Benson, T.E., Kim, D.H. & Lees, W.J. (1996). The versatility of phosphoenolpyruvate and its vinyl ether products in biosynthesis. *Chem. & Biol.* **3**, 83–91.
9. Steinrücken, H.C. & Amrhein, N. (1980). The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimic acid 3-phosphate synthase. *Biochem. Biophys. Res. Commun.* **94**, 1207–1212.
10. Wanke, C., Falchetto, R. & Amrhein, N. (1992). The UDP-*N*-acetylglucosamine 1-carboxyvinyl-transferase of *Enterobacter cloacae*. Molecular cloning, sequencing of the gene and overexpression of the enzyme. *FEBS Lett.* **301**, 271–276.
11. Anderson, K.S. & Johnson, K.A. (1990). Kinetic and structural analysis of enzyme intermediates: lessons from the EPSP synthase. *Chem. Rev.* **90**, 1131–1149.
12. Appleyard, R.J., Shuttleworth, W.A. & Evans, J.N.S. (1994). Time-resolved solid-state NMR of EPSP synthase. *Biochemistry* **33**, 6812–6821.
13. Kim, D.H., Lees, W.J., Kempsell, K.E., Lane, W.S., Duncan, K. & Walsh, C.T. (1996). Characterization of a Cys115 to Asp substitution in the *Escherichia coli* cell wall biosynthetic enzyme UDP-GlcNAc enolpyruvyl transferase (MurA) that confers resistance to inactivation by the antibiotic fosfomycin. *Biochemistry* **35**, 4923–4928.
14. Steinrücken, H.C. & Amrhein, N. (1984). 5-Enolpyruvylshikimate-3-phosphate synthase of *Klebsiella pneumoniae*. 2. Inhibition by glyphosate [N-(phosphonomethyl)glycine]. *Eur. J. Biochem.* **143**, 351–357.
15. Stallings, W.C., *et al*., & Kishore, G.M. (1991). Structure and topological symmetry of the glyphosate target 5-enolpyruvylshikimate-3-phosphate synthase: a distinctive protein fold. *Proc. Natl. Acad. Sci. USA* **88**, 5046–5050.
16. Sack, S., Dauter, Z., Wanke, C., Amrhein, N., Mandelkow, E. & Schönbrunn, E. (1996). Crystallization and preliminary X-ray diffraction analysis of UDP-*N*-acetylglucosamine enolpyruvyltransferase of *Enterobacter cloacae*. *J. Struct. Biol.* **117**, 73–76.
17. Marquardt, J.L., Siegele, D.A., Kolter, R. & Walsh, C.T. (1992). Cloning and sequencing of *Escherichia coli murZ* and purification of its product, a UDP-*N*-acetylglucosamine enolpyruvyl transferase. *J. Bacteriol.* **174**, 5748–5752.
18. Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
19. Padgette, S.R., *et al*., & Fraley, R.T. (1996). New weed control opportunities: development of soybeans with a Roundup Ready™ gene. In *Herbicide Resistant Crops: Agricultural, Environmental, Economic, Regulatory, and Technical Aspects*. (Duke, S.O., ed), pp. 53–84, CRC Lewis Publishers, Boca Raton, NY.
20. Gerstein, M., Lesk, A.M. & Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry* **33**, 6739–6749.
21. Schönbrunn, E. & Koch, M.H.J. (1995). Small-angle X-ray scattering study on enolpyruvyltransferase. In *HASYLAB Annual Report*. pp. 853–854, DESY, Hamburg.
22. Svergun, D., Barberato, C. & Koch, M.H.J. (1995). CRYSOL-a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* **28**, 768–773.
23. Cassidy, P.J. & Kahan, F.M. (1973). A stable enzyme-phosphoenolpyruvate intermediate in the synthesis of uridine-5′-diphospho-*N*-acetyl-2-amino-2-deoxyglucose 3-O-enolpyruvyl ether. *Biochemistry* **12**, 1364–1374.
24. Otwinowski, Z. (1993). In *Data collection and Processing, Proceedings of the CCP4 Study Weekend.* (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
25. Furey, W. & Swaminathan, S. (1996). PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules. In *Macromolecular Crystallography Methods in Enzymology.* (Carter, C. & Sweet, R., eds), Academic Press, Orlando, FL, in press.
26. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
27. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A* **50**, 157–163.
28. Collaborative Computational Project No.4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–767.
29. Brünger, A.T. (1992). *X-PLOR Manual Version* 3.1: A system for X-ray crystallography and NMR. Yale University Press, New Haven, CT, USA.
30. Brünger, A.T., Krukowski, A. & Erickson, J. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Cryst. A* **46**, 585–593.
31. Brünger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.
32. Lamzin, V.S. & Wilson, K.S. (1993). Automated refinement of protein models. *Acta Cryst. D* **49**, 129–147.
33. Konnert, J.H. & Hendrickson, W.A. (1980). A restrained-parameter thermal factor refinement procedure. *Acta Cryst. A* **36**, 344–350.
34. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.
35. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.
36. Kleywegt, G.J. & Jones, T.A. (1994). A super position. *ESF/CCP4 Newsletter* **31**, 9–14.